

Coupled Wavepackets for Non-Adiabatic Molecular Dynamics: A Generalization of Gaussian Wavepacket Dynamics to Multiple Potential Energy Surfaces

Alexander White, Sergei Tretiak and Dmitry Mozysky

Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

(Dated: February 12, 2016)

Accurate simulation of the non-adiabatic dynamics of molecules in excited electronic states is key to understanding molecular photo-physical processes. Here we present a novel method, based on a semiclassical approximation, that is as efficient as the commonly used mean field Ehrenfest or *ad hoc* Surface Hopping methods and properly accounts for interference and decoherence effects. This novel method is an extension of Heller's Thawed Gaussian wavepacket dynamics that includes coupling between potential energy surfaces. The accuracy of the method can be systematically improved.

First principles-based molecular dynamics (MD) is becoming an important tool for understanding properties of complex molecular systems.[1–3] Unfortunately, the cost of exact dynamics, by direct calculation of the time-dependent Schrödinger equation (TDSE), scales exponentially with the dimensionality (*i.e.* number of atoms) of the system.[4–10] Thus, for large systems one often approximates that the nuclei of a molecule propagate via classical equations of motion and calculates forces (due to Coulombic interaction) via quantum chemistry methods. In doing so one typically relies on the *Born-Oppenheimer* approximation, where electrons remain in the same electronic quantum state $|n(\mathbf{x})\rangle$ with energy, $E^{(n)}(\mathbf{x})$, that parametrically depends on nuclear coordinates, $\mathbf{x} = (x_1, \dots, x_N)^T$. [11] Thus, the nuclei propagate on a single potential energy surface (PES). For molecules in the ground electronic state, and at low temperatures, this situation often holds due to a sufficiently wide gap between the PES of the ground and excited electronic states. However, for certain nuclear configurations, common when the molecule is in an excited electronic state due to absorption of energy (*e.g.* a photon), the gaps can become small or even vanish. In these regions, where the nuclear-electronic coupling is the same order as the energy gap, non-adiabatic behavior is expected.[12] This creates a superposition of electronic states, with different forces acting on the nuclei.

Since the full TDSE is numerically intractable for high dimensions, approximations for the non-adiabatic molecular dynamics (NAMD) must be made. The simplest approximation is to average over the electronic degree of freedom (DOF), a mean-field approximation, to determine the force on the nuclei.[13, 14] This is known as the Ehrenfest approximation. Like any mean-field approximation, it breaks down when there is non-negligible correlation between the dynamical DOF (the nuclear) and the averaged DOF (the electronic), *i.e.* if the components of the nuclear wavefunction separate depending on which PES they propagate. In an attempt to correct for this problem, while maintaining efficiency and simplicity, Tully proposed the surface hopping method,[15] most commonly used with the fewest-switching proce-

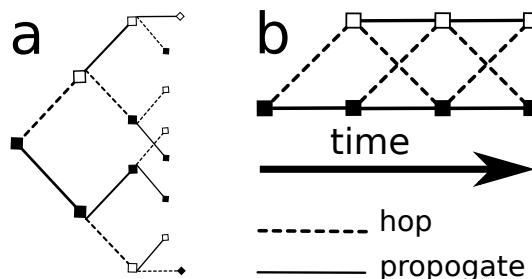


FIG. 1. a- Branching Tree Solution to time dependent Schrödinger equation (sampled by Monte-Carlo). b- Coupled wavepackets for non-adiabatic molecular dynamics (CW-NAMD) approximation to the Branching Tree. c- CW-NAMD approximation with coarse branching.

dure (FSSH).[16] In this method a swarm of classical trajectories propagate on an initial PES, with a finite probability to hop to a coupled PES in regions of non-adiabatic coupling. This method is *ad-hoc*, and is only strictly accurate in the same limit as the Ehrenfest approximation.[17, 18] This incomplete treatment of the nuclear-electron correlation has two well known symptoms: the interference problem, where the incorrect phase of the nuclear wavefunction leads to incorrect levels of constructive/deconstructive interference, and the decoherence problem, where the separation of the nuclear wavefunction is improperly accounted for. Both problems were pointed out by Tully himself.[16] These two methods, Ehrenfest and FSSH, are by far the most commonly used in the simulation of NAMD.[19–31] Many attempts have been made to improve upon the basic foundation of these two methods.[17, 18, 32–40] More sophisticated and accurate mixed quantum-classical and semi-classical methods, which are not *ad-hoc*, are typically applied only in small, or reduced, systems due to inefficiency and/or complexity.[41–55]

An ideal NAMD method would have certain properties. It must (1) be based on localized dynamics, *i.e.* based on real-space trajectories, (2) use only local parameters easily calculated from common electronic structure methods, *i.e.* PES and electronic wavefunction, (3)

require no empirical or *ad-hoc* treatments, (4) include proper treatment of electron-nuclear coupling, (5) be at least as efficient as surface hopping, and (6) be systematically improvable. To build a new NAMD method, which satisfies the first and second conditions, one must start from a sound foundation for these real-spaced trajectories. The closest analog to a classical particle, and thus real local trajectory, for a quantum system is a localized wave packet, or superposition of wave packets. [56]

The use of complex multi-dimensional Gaussian wave packets (GWP):

$$g(\mathbf{x}; \mathbf{x}_0, \mathbf{p}_0, \hat{\alpha}_0) = e^{\frac{i}{\hbar}[\gamma_0 + \mathbf{p}_0^T(\mathbf{x} - \mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0)^T \hat{\alpha}_0(\mathbf{x} - \mathbf{x}_0)]}, \quad (1)$$

as approximations, or basis sets, for nuclei wavefunctions is well studied for semiclassical dynamics on a single potential energy surfaces.[57–64] In 1975, Heller derived the equations of motion for the four parameters (position \mathbf{x}_0 , momentum \mathbf{p}_0 , complex width matrix $\hat{\alpha}_0$, complex phase γ_0) of the GWP, assuming the PES is locally quadratic around \mathbf{x}_0 , the *Thawed Gaussian* approximation (TGA). The key of this method is that the phase-space center of the wavepacket moves by classical mechanics. That classical point is “dressed” in the semiclassical width and phase. [57, 65]

In the adiabatic limit, the dynamics can be formally described in the framework of quantum mechanical description of the nuclei, $\Psi(\mathbf{x}, t) = e^{iH(\mathbf{x})t}\Psi(\mathbf{x}, 0)$ (here and in the following we set $\hbar = 1$ unless stated otherwise),

$$H(\mathbf{x}) = -\sum_i \frac{1}{2m_i} \frac{\partial^2}{\partial x_i^2} + V(\mathbf{x}), \quad (2)$$

where $H(\mathbf{x})$ is the Hamiltonian of the system. The potential $V(\mathbf{x})$ is a parametric function of geometry \mathbf{x} , m is the nuclear mass, and $\Psi(\mathbf{x}, t)$ is the nuclear wavefunction. TGA can be alternatively derived by splitting the evolution operator e^{-iHt} into slices $e^{-iH\epsilon} e^{-iH\epsilon} \dots$ with an infinitesimally small time step ϵ . [65] If for a single time slice one expands $e^{iH(\mathbf{x})\epsilon}$ to first order in ϵ , applies the same approximation as Heller, and re-exponentiates, one recovers a new Gaussian with parameters shifted by one time step using Heller’s equations of motion.[57, 65] We seek to follow similar steps to generalize TGA for multiple electronic states.

The non-adiabatic dynamics can similarly be obtained from a quantum mechanical description of the nuclei, $|\Psi(\mathbf{x}, t)\rangle = e^{i\hat{H}(\mathbf{x})t}|\Psi(\mathbf{x}, 0)\rangle$. Now the nuclei’s potential energy operator $\hat{V}(\mathbf{x})$ and the wavefunction $|\Psi(\mathbf{x}, t)\rangle$ are $M \times M$ hermitian matrix and M component vector respectively, where M is the number of relevant electronic states. For simplicity we will consider a situation with two levels crossing, i.e., with $M = 2$. This is the most common situation, typically more complex problems with multiple PESs and crossings can be modeled as consecutive transitions through well separated regions of coupling between two locally adjacent

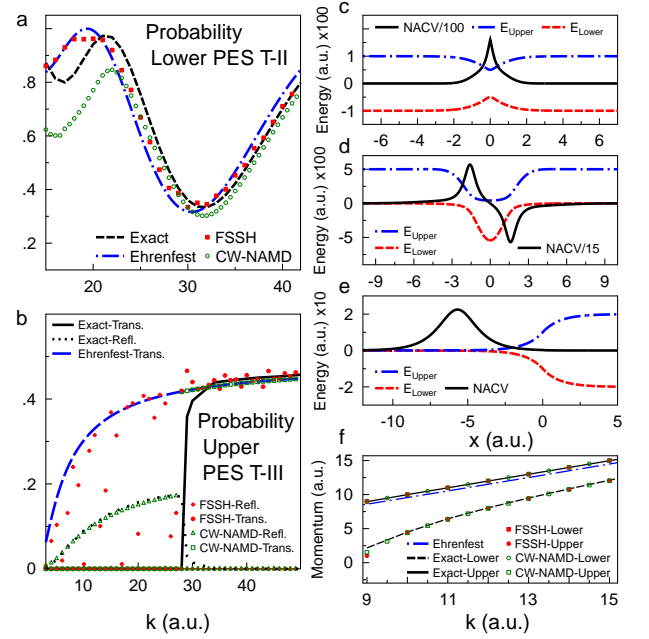


FIG. 2. a (b)- Scattering probabilities Tully II (III) problems on the Lower (Upper) surface for different initial wave vectors k . Exact solution, FSSH (2000 trajectories), Ehrenfest and CW-NAMD are compared. Initial wavepacket position $x_{\text{initial}} = -10$ a.u.. Initial width, $\alpha_{\text{initial}} = ik^2/400$ for all. c (d,e)- Potential Energy Surfaces (E) and Non-Adiabatic Coupling Vectors (NACV) for Tully I (II, III). g- Average momentum for each surface after scattering (Tully I). Exact solution, FSSH, Ehrenfest and RG-WP are compared.

PESs. Furthermore, an extension to the $M > 2$ situation is straightforward. We assume that the initial state is a single Gaussian localized on the first PES, $|\Psi(\mathbf{x}, 0)\rangle = N_0^{(1)} g^{(1)}(\mathbf{x}; \mathbf{x}_0^{(1)}, \mathbf{p}_0^{(1)}, \hat{\alpha}_0^{(1)}) |1[\mathbf{x}_0^{(1)}]\rangle$, where $|1[\mathbf{x}_0^{(1)}]\rangle$ an eigenstate of $\hat{V}(\mathbf{x})$ corresponding to the first PES evaluated at $\mathbf{x}_0^{(1)}$ and $N_0^{(1)}$ is the real amplitude of the otherwise normalized state $|\Psi(\mathbf{x}, 0)\rangle$. Here and in the following the superscripts indicate the electronic state or PES. Non-Gaussian states can be treated as linear superpositions of finite number of Gaussians due to the linearity of the TDSE.

We again split the evolution operator $e^{-i\hat{H}t}$ into slices $e^{-i\hat{H}\epsilon} e^{-i\hat{H}\epsilon} \dots$ and now introduce a basis resolution $\sum_{i=1,2} |i[\mathbf{x}_1^{(1)}]\rangle \langle i[\mathbf{x}_1^{(1)}]|$ between the first and the second slices (the subscripts here and below indicate the time steps). The point $\mathbf{x}_1^{(1)}$ is the location of the classical trajectory to be specified below. The use of the local electronic basis function is the first deviation from previously derived path-integral GWP dynamics.[66, 67] Physically introduction of the basis resolution corresponds to projecting the wavepacket on the new, slightly shifted basis of the eigenstates of \hat{V} at the new average position of the wavepacket at time ϵ . The new wavepacket will mostly

remain in the electronic state $|1[\mathbf{x}_1^{(1)}]\rangle$ with a small ($\propto \epsilon$) transfer to $|2[\mathbf{x}_1^{(1)}]\rangle$. After some calculation one gets [65]

$$|\Psi(\mathbf{x}, \epsilon)\rangle = N_1^{(1)} g^{(1)}(\mathbf{x}) |1[\mathbf{x}_1^{(1)}]\rangle + \epsilon N_1^{(2)} g^{(2)}(\mathbf{x}) |2[\mathbf{x}_1^{(1)}]\rangle. \quad (3)$$

The change in the wavepacket $g^{(1)}(\mathbf{x})$ in Eq. 3 (i.e., after a single time step) is infinitesimal with the same form as the Heller GWP dynamics, leading to equations of motion for the multistate case:

$$\begin{aligned} \dot{\mathbf{x}}_0^{(1)} &= \mathbf{p}_0^{(1)} \hat{m}^{-1} \\ \dot{\mathbf{p}}_0^{(1)} &= -\langle 1[\mathbf{x}_0^{(1)}] | \partial_{\mathbf{x}} \hat{V}(\mathbf{x}_0) | 1[\mathbf{x}_0^{(1)}] \rangle, \\ \dot{\hat{\alpha}}_0^{(1)} &= -2\hat{\alpha}_0^{(1)} \hat{m}^{-1} \hat{\alpha}_0^{(1)} - \frac{1}{2} \langle 1[\mathbf{x}_0^{(1)}] | \partial_{\mathbf{x}}^2 \hat{V}(\mathbf{x}_0) | 1[\mathbf{x}_0^{(1)}] \rangle, \\ \dot{\gamma}_0 &= i\hbar \text{Tr}[\hat{\alpha}_0 \hat{m}^{-1}] + \frac{1}{2} \mathbf{p}_0 \hat{m}^{-1} \mathbf{p}_0 - \langle 1[\mathbf{x}_0^{(1)}] | \hat{V}(\mathbf{x}_0) | 1[\mathbf{x}_0^{(1)}] \rangle. \end{aligned} \quad (4)$$

The weight, $N_1^{(1)} = N_0^{(1)}$, is unchanged.

The wavepacket $g^{(2)}(\mathbf{x})$ “hopped” to PES 2. It has the same classical position as the original wavepacket, $\mathbf{x}_1^{(2)} = \mathbf{x}_0^{(1)}$, but different momentum: $\mathbf{p}_1^{(2)}$ is such that $\mathbf{p}_1^{(2)} - \mathbf{p}_0^{(1)}$ is parallel to the non-adiabatic coupling vector $\mathbf{d}_{12}(\mathbf{x}_0^{(1)}) = \langle 2[\mathbf{x}_0^{(1)}] | \partial_{\mathbf{x}} | 1[\mathbf{x}_0^{(1)}] \rangle$, and its absolute value satisfies the energy conservation condition, $\sum_{\alpha=1}^N [(p_{1\alpha}^{(2)})^2 - (p_{0\alpha}^{(1)})^2] / (2m_\alpha) = E^{(1)}(\mathbf{x}_0^{(1)}) - E^{(2)}(\mathbf{x}_0^{(1)})$ [65, 68]. This rescaled momentum is a direct consequence of the projection onto local electronic basis functions. The parameters $\alpha^{(2)}$ and $N_1^{(2)}$ are related to the coefficients of $g^{(1)}(\mathbf{x})$ as

$$\begin{aligned} \alpha_1^{(2)} &= \alpha_0^{(1)} + \frac{1}{2} \frac{\langle 2[\mathbf{x}_0^{(1)}] | \partial_{\mathbf{x}}^2 \hat{V}(\mathbf{x}_0) | 1[\mathbf{x}_0^{(1)}] \rangle}{[\mathbf{d}^{(12)}(\mathbf{x}_0^{(1)}) \cdot \bar{\mathbf{v}}_1^{(1)}]}, \\ N_1^{(12)} &= N_0^{(1)} \mathbf{d}^{(12)}(\mathbf{x}_0^{(1)}) \cdot \bar{\mathbf{v}}_1^{(1)} \exp \left[\frac{\mathbf{d}^{(12)}(\mathbf{x}_0^{(1)}) \cdot \Delta \mathbf{v}_0^{(1)}}{\mathbf{d}^{(12)}(\mathbf{x}_0^{(1)}) \cdot \bar{\mathbf{v}}_1^{(1)}} \right], \end{aligned} \quad (5)$$

where $v_{0\alpha}^{(1)} = p_{0\alpha}^{(1)} / m_\alpha$, $\bar{\mathbf{v}}_1^{(1)} = (\mathbf{v}_0^{(1)} + \mathbf{v}_1^{(2)}) / 2$ and $\Delta \mathbf{v}_0^{(1)} = (\mathbf{v}_0^{(1)} - \mathbf{v}_1^{(2)}) / 2$. Note that the parameters of the spawned wavepacket, e.g. Eqs. 5, change discontinuously at the moment of the hop. In practice we only keep the linear term in the expansion of $\hat{V}(\mathbf{x})$, since for realistic systems calculation of the quadratic term can be very costly.

At the next time step each of the wavepackets propagates and spawns again, according to Eqs. 3 - 5 (with a replacement $1 \rightarrow 2$ for the wavepacket on the second PES). After each time step the total number of the wavepackets doubles. Such process can be viewed as branching on a tree, shown in Fig. 1-a. This branching tree can be evaluated by a Monte-Carlo approach [49, 51, 69] which becomes too computationally expensive in systems with multiple level crossings.

Here we propose a new approach based on the wavepacket reconstruction after each spawning event.

The approach is schematically shown in Fig. 1-b. That is, after two time steps, described in Eqs. (3 - 5), one creates two wavepackets, on each PES, which will give rise to four more, etc. We note, however, that if each pair of the wavepackets on the same surface has close coordinates and momenta, one can replace each pair by a single GWP, with slightly shifted parameters. We parameterize the new Gaussian by calculating the expectation values of \hat{x} , \hat{p} , \hat{x}^2 , \hat{p}^2 of the superposition. $\langle \hat{x} \rangle$ and $\langle \hat{p} \rangle$ are taken as the position and momentum of the RG wavepacket, while $\langle \hat{x}^2 \rangle$ and $\langle \hat{p}^2 \rangle$ directly give the new complex width. The new phase and weight, γ and N , are determined by maximizing the overlap of the new wavepacket with the superposition, under the constraint that $|N|^2$ is the same as the density of the superposition.[65] Approximations are made in order to decouple the calculation of $\langle \hat{x} \rangle$ and $\langle \hat{p} \rangle$ from the explicit form of the wavefunction, i.e. $\hat{\alpha}$. [65] Thus, as with Heller’s equations, the trajectories of the GWPs remains independent of the phase and width. At the next step the procedure is repeated, again we have only two GWP and so on. The process repeats until the overlap, O_{12} , between the gaussians within each pair becomes intolerable, $O_{12} < O_{min}$. At this point, or if the non-adiabatic coupling drops below its own threshold, the “coupling” between the GWPs stops and each is treated independently, thus new branching is allowed. This coarse branching is schematically shown in Fig. 1-c. This approximation significantly reduces or eliminates the exponential growth of the number of wavepackets. We call this approximation coupled wavepackets for non-adiabatic molecular dynamics (CW-NAMD).

As the two wave packets separate in position space, their electronic bases will become non-orthogonal. Formally this must be taken into account by considering the required basis rotations when reconstruction occurs. These rotations lead to a correction, but it is small and does not affect the results presented in this letter.[65]

The CW-NAMD method is similar in spirit to the *ab initio* multiple spawning (AIMS) method developed by Martinez *et. al.* [70–72] Both involve approximate solution to an infinitely branching tree, of GWPs. However, in practice AIMS is usually based on independent, frozen GWP, which are non-interfering. CW-NAMD uses thawed GWPs and considers the full superposition of GWPs. For AIMS the “spawning” procedure is based on well-reasoned but empirical considerations,[73] and is only truly *ab initio* in the limit of infinite spawning, so called Full AIMS. The branching procedure in CW-NAMD has a simple numerical control parameter, O_{min} .

Figure 2 shows scattering probabilities for the standard Tully test problems II and III, Figure 2-d, e. These problems are frequently used to test new methods of non-adiabatic dynamics because they specifically probe the interference (Tully II) and decoherence (Tully III) questions directly. We compare the CW-NAMD results, with $O_{min} = 0$, to the standard fewest switching surface hop-

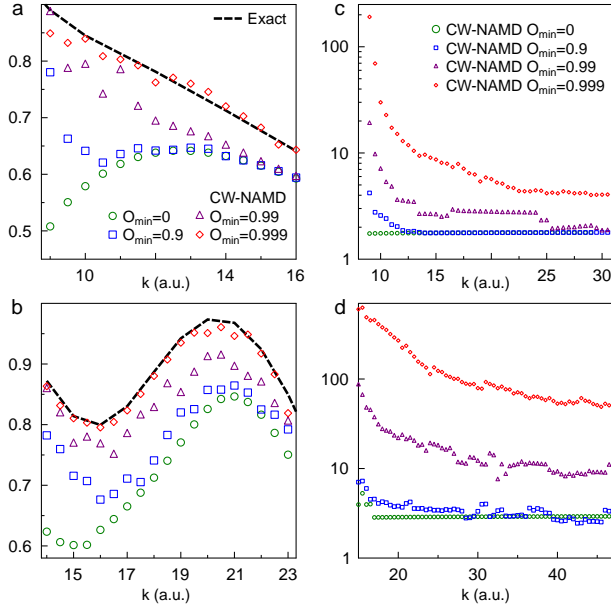


FIG. 3. (a) (b)- Comparison of low momentum transmission probabilities, on lower surface, with different values of O_{min} , compared to exact solution, for Tully I (II). Initial conditions as in Fig 2 (except $x_{initial} = -5$ a.u. for Tully I). (c,d)- Number of “effective” trajectories for Tully I (II) calculation with different values of O_{min} . Dynamics are run for a total time of $\frac{25,000}{k}$ ($\frac{40,000}{k}$) a.u. for Tully I (II).

ping (FSSH) and the mean-field Ehrenfest method as well as direct calculation of the time-dependent Schrödinger equation. When branching does not occur, the computational cost of the CW-NAMD method is similar to Ehrenfest (i.e. there is one force calculation per surface per time point), and is much lower than surface hopping. Figure 2 a,b demonstrates that for sufficiently high momentum the CW-NAMD method produces the correct scattering results. The CW-NAMD does not suffer from the interference or decoherence errors of Ehrenfest or FSSH. This can be observed by comparing the position of the peaks of the Stueckelberg oscillations [74] in Figure 2-a and the lack of false oscillations in the reflected probabilities in Figure 2-b. Unlike Ehrenfest, CW-NAMD produces the correct momenta and positions of the wave packets on the upper and lower surface (see Figure 2-f). However at low momenta, the total scattering probability is not conserved and may be poorly estimated. This is evident in both Tully-II (see Figure 2-a, 3-b) and Tully-I (3-a). This can be corrected by allowing the coupled GWPs to branch, i.e. set $O_{min} > 0$.

We compare the low momentum results for Tully I (II) with different values of O_{min} in Figure 3 a (b). The difference between exact and CW-NAMD solutions is systematically improved by increasing O_{min} . The increased cost can be seen in Figure 3 c (d). In direct dynam-

ics simulations the bottleneck is typically the calculation of the PES gradients (forces). Trajectory methods like FSSH, require one force calculation per time step per trajectory. Thus we define an “effective” number of trajectories, by determining the total number of force calculations (summed over all branches) divided by the total number of time steps for the simulation, to compare the cost of a branching scheme to that of a trajectory based methods (i.e. FSSH). We see a growth of the number of trajectories required with increased O_{min} , However the cost of CW-NAMD is still lower than the 2000 trajectories used to calculate the FSSH result (Fig. 2 b). In the limit $O_{min} = 1$ we recover the full branching tree (Fig. 1 a). Lower values of O_{min} result in a coarse-grained tree (Fig. 1 c). To prevent overgrowth of the tree, we place hard-limits on the spawning rate and utilize pruning procedures to discard irrelevant branches. [65]

In conclusion, the new CW-NAMD method is a highly efficient and accurate method of simulating non-adiabatic dynamics applicable to realistic molecular systems. CW-NAMD consistently accounts for decoherence and interference between different dynamical pathways. It can be as efficient as the Ehrenfest method in the high momentum limit, moreover it accurately describes the dynamics of branching wave packets. In the low momentum limit the method can be systematically improved by increase the rate of allowed branching via the user controlled accuracy threshold, O_{min} . Combined with filtering of insignificant branches, the method is more accurate and more efficient than the standard FSSH. In our test problems we observe numerical cost of CW-NAMD ranging from about $2(M)$ to 500 trajectories depending on initial momentum and desired accuracy. This needs to be compared with the number of effective trajectories in other methods: $2(M)$ (Ehrenfest), $(2 - 5) \times 10^3$ (FSSH), $(2 - 10) \times 10^4$ (Monte-Carlo approaches). The development of CW-NAMD opens new avenues for future research: more advance branching criterion, manipulation of the electronic bases, optimization of the reconstruction and branch pruning procedures, and application to molecular systems of increasing size.

We acknowledge support of the U.S. Department of Energy through the Los Alamos National Laboratory (LANL) LDRD Program. LANL is operated by Los Alamos National Security, LLC, for the National Nuclear Security Administration of the U.S. Department of Energy under Contract No. DE-AC52-06NA25396. We also acknowledge the LANL Institutional Computing (IC) Program provided computational resources. This work was supported in part by the Center for Nonlinear Studies (CNLS) and the Center for Integrated Nanotechnology (CINT) at LANL.

-
- [1] *Ab Initio Molecular Dynamics* (Cambridge University Press).
- [2] M. E. Tuckerman, *Journal of Physics: Condensed Matter* **14**, R1297 (2002).
- [3] R. Car and M. Parrinello, *Phys. Rev. Lett.* **55**, 2471 (1985).
- [4] M. Feit, J. Fleck, and A. Steiger, *Journal of Computational Physics* **47**, 412 (1982).
- [5] H. TalEzer and R. Kosloff, *The Journal of Chemical Physics* **81**, 3967 (1984).
- [6] T. J. Park and J. C. Light, *The Journal of Chemical Physics* **85**, 5870 (1986).
- [7] R. Kosloff, *Annual Review of Physical Chemistry* **45**, 145 (1994).
- [8] D. Neuhauser, *The Journal of Chemical Physics* **100**, 9272 (1994).
- [9] C. Xie, J. Ma, X. Zhu, D. H. Zhang, D. R. Yarkony, D. Xie, and H. Guo, *The Journal of Physical Chemistry Letters* **5**, 1055 (2014).
- [10] B. Jiang and H. Guo, *Phys. Rev. Lett.* **114**, 166101 (2015).
- [11] M. Born and R. Oppenheimer, *Annalen der Physik* **389**, 457 (1927).
- [12] H. S. W. Massey, *Reports on Progress in Physics* **12**, 248 (1949).
- [13] P. Ehrenfest, *Zeitschrift Physik* **45**, 455 (1927).
- [14] S.-I. Sawada, A. Nitzan, and H. Metiu, *Phys. Rev. B* **32**, 851 (1985).
- [15] J. C. Tully and R. K. Preston, *The Journal of Chemical Physics* **55**, 562 (1971).
- [16] J. C. Tully, *The Journal of Chemical Physics* **93**, 1061 (1990).
- [17] J. E. Subotnik, *The Journal of Chemical Physics* **132**, 134112 (2010).
- [18] J. E. Subotnik, *The Journal of Physical Chemistry A* **115**, 12083 (2011).
- [19] S. V. Kilina, D. S. Kilin, and O. V. Prezhdo, *ACS Nano* **3**, 93 (2009).
- [20] K. Saita and D. V. Shalashilin, *The Journal of Chemical Physics* **137**, 22 (2012).
- [21] T. Nelson, S. Fernandez-Alberti, A. E. Roitberg, and S. Tretiak, *Accounts of Chemical Research* **47**, 1155 (2014).
- [22] A. V. Akimov and O. V. Prezhdo, *Journal of Chemical Theory and Computation* **10**, 789 (2014).
- [23] L. Wang, R. Long, and O. V. Prezhdo, *Annual Review of Physical Chemistry* **66**, 549 (2015).
- [24] S. Choi, R. Onofrio, and B. Sundaram, *Phys. Rev. E* **92**, 042907 (2015).
- [25] P. Goyal, C. A. Schwerdtfeger, A. V. Soudackov, and S. Hammes-Schiffer, *The Journal of Physical Chemistry B* **119**, 2758 (2015).
- [26] A. S. Petit and J. E. Subotnik, *Journal of Chemical Theory and Computation* **11**, 4328 (2015).
- [27] L. Du and Z. Lan, *Journal of Chemical Theory and Computation* **11**, 1360 (2015).
- [28] W. Ouyang, W. Dou, and J. E. Subotnik, *The Journal of Chemical Physics* **142**, 084109 (2015).
- [29] W. Dou, A. Nitzan, and J. E. Subotnik, *The Journal of Chemical Physics* **142**, 084110 (2015).
- [30] W. Dou, A. Nitzan, and J. E. Subotnik, *The Journal of Chemical Physics* **142**, 234106 (2015).
- [31] M. Galperin and A. Nitzan, *The Journal of Physical Chemistry Letters* **6**, 4898 (2015).
- [32] E. R. Bittner and P. J. Rossky, *The Journal of Chemical Physics* **103**, 8130 (1995).
- [33] C. Zhu, S. Nangia, A. W. Jasper, and D. G. Truhlar, *The Journal of Chemical Physics* **121**, 7658 (2004).
- [34] M. J. Bedard-Hearn, R. E. Larsen, and B. J. Schwartz, *The Journal of Chemical Physics* **123**, 234106 (2005), <http://dx.doi.org/10.1063/jcp.123.234106>.
- [35] T. Nelson, S. Fernandez-Alberti, A. E. Roitberg, and S. Tretiak, *The Journal of Chemical Physics* **138**, 224111 (2013).
- [36] A. V. Akimov and O. V. Prezhdo, *Phys. Rev. Lett.* **113**, 153003 (2014).
- [37] L. Wang and O. V. Prezhdo, *The Journal of Physical Chemistry Letters* **5**, 713 (2014).
- [38] A. E. Sifain, L. Wang, and O. V. Prezhdo, *The Journal of Chemical Physics* **142**, 224102 (2015).
- [39] L. Wang, A. E. Sifain, and O. V. Prezhdo, *The Journal of Physical Chemistry Letters* **6**, 3827 (2015).
- [40] D. J. Trivedi and O. V. Prezhdo, *The Journal of Physical Chemistry A* **119**, 8846 (2015).
- [41] H. Meyera and W. H. Miller, *The Journal of Chemical Physics* **70**, 3214 (1979).
- [42] M. Thoss and G. Stock, *Phys. Rev. A* **59**, 64 (1999).
- [43] R. Kapral, *Annual Review of Physical Chemistry* **57**, 129 (2006).
- [44] X. Chen and V. S. Batista, *The Journal of Chemical Physics* **125**, 124313 (2006).
- [45] W. H. Miller, *The Journal of Physical Chemistry A* **113**, 1405 (2009).
- [46] M. F. Herman, *The Journal of Physical Chemistry A* **109**, 9196 (2005).
- [47] V. A. Rassolov and S. Garashchuk, *Phys. Rev. A* **71**, 032511 (2005).
- [48] N. Zamstein and D. J. Tannor, *The Journal of Chemical Physics* **137**, 22 (2012).
- [49] V. N. Gorshkov, S. Tretiak, and D. Mozyrsky, *Nature Communications* **4**, 2144 (2013).
- [50] A. R. Menzelev, F. Bell, and T. F. Miller, *The Journal of Chemical Physics* **140**, 064103 (2014).
- [51] A. J. White, V. N. Gorshkov, R. Wang, S. Tretiak, and D. Mozyrsky, *The Journal of Chemical Physics* **141**, 184101 (2014).
- [52] N. Makri, *International Journal of Quantum Chemistry* **115**, 1209 (2015).
- [53] S. K. Min, F. Agostini, and E. K. U. Gross, *Phys. Rev. Lett.* **115**, 073001 (2015).
- [54] W. C. Pfalzgraff, A. Kelly, and T. E. Markland, *The Journal of Physical Chemistry Letters* **6**, 4743 (2015).
- [55] C. C. Martens, *The Journal of Chemical Physics* **143**, 141101 (2015).
- [56] *The Principles of Quantum Mechanics* (Oxford University Press, 1958).
- [57] E. J. Heller, *The Journal of Chemical Physics* **62**, 1544 (1975).
- [58] E. J. Heller, *The Journal of Chemical Physics* **75**, 2923 (1981).
- [59] D. Huber and E. J. Heller, *The Journal of Chemical Physics* **87**, 3302 (1987).
- [60] R. D. Coalson and M. Karplus, *The Journal of Chemical Physics* **93**, 3919 (1990).
- [61] A. K. Pattanayak and W. C. Schieve, *Phys. Rev. E* **50**, 3601 (1994).
- [62] M. F. Herman and E. Kluk, *Chemical Physics* **91**, 27 (1984).

- [63] E. Kluk, M. F. Herman, and H. L. Davis, *The Journal of Chemical Physics* **84**, 326 (1986).
- [64] B. Gu and S. Garashchuk, *The Journal of Physical Chemistry A* **0**, Article ASAP (0).
- [65] See Supplemental Material at [URL will be inserted by publisher] for details.
- [66] S. Kremp, M. Winterstetter, H. Plhn, and W. Domcke, *The Journal of Chemical Physics* **100**, 926 (1994).
- [67] R. D. Coalson, *The Journal of Physical Chemistry* **100**, 7896 (1996).
- [68] M. F. Herman, *The Journal of Chemical Physics* **81**, 754 (1984).
- [69] A. J. White, V. N. Gorshkov, S. Tretiak, and D. Mozyrsky, *The Journal of Chemical Physics* **143**, 014115 (2015).
- [70] T. J. Martinez, M. Ben-Nun, and R. D. Levine, *The Journal of Physical Chemistry* **100**, 7884 (1996).
- [71] M. Ben-Nun, J. Quenneville, and T. J. Martinez, *The Journal of Physical Chemistry A* **104**, 5161 (2000).
- [72] T. J. Martinez, *Accounts of Chemical Research* **39**, 119 (2006).
- [73] B. G. Levine, J. D. Coe, A. M. Virshup, and T. J. Martinez, *Chemical Physics* **347**, 3 (2008).
- [74] E. C. G. Stueckelberg, *Helvetica Physica Acta* **5**, 370 (1932).

Supplemental Materials: Coupled Wavepackets for Non-Adiabatic Molecular Dynamics: A Generalization of Gaussian Wavepacket Dynamics to Multiple Potential Energy Surfaces

Alexander White, Sergei Tretiak, and Dmitry Mozyrsky

Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA

I. HELLER'S ADIABATIC GAUSSIAN WAVEPACKET DYNAMICS FROM PATH INTEGRAL APPROACH

Here we re-derive the *Thawed Gaussian* approximation, originally derived by Heller, [1] based on a time slicing procedure. This will provide the foundation from which we will extend the derivation to multiple coupled potential energy surfaces. The wavefunction of the system at time t is given by the time dependent Schrödinger Equation and the initial wavefunction:

$$\Psi(\mathbf{x}, t) = e^{-i\varepsilon H(\mathbf{x})} e^{-i\varepsilon H(\mathbf{x})} \dots e^{-i\varepsilon H(\mathbf{x})} e^{-i\varepsilon H(\mathbf{x})} \Psi(\mathbf{x}, 0) . \quad (1)$$

Here we have defined $\varepsilon \equiv \frac{t}{M\hbar}$, where M is a large number. For a single time step, we expand to first order in ε :

$$\begin{aligned} \Psi(\mathbf{x}, t) &= \dots \{1 - i\varepsilon K(\mathbf{x}) - i\varepsilon V(\mathbf{x})\} \Psi(\mathbf{x}, 0) = \\ &\dots \{1 - i\varepsilon (\sum_i \frac{-\hbar^2}{2m_i} \frac{\partial^2}{\partial x_i^2} + V(\mathbf{x}))\} \Psi(\mathbf{x}, 0) \\ &\approx \dots \Psi(\mathbf{x}, \hbar\varepsilon) . \end{aligned} \quad (2)$$

Using similar notation to Heller the initial wavepacket is given by:[1]

$$\Psi(\mathbf{x}, 0) = \exp[\frac{i}{\hbar} \{ \gamma_0 + \mathbf{p}_0^T (\mathbf{x} - \mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0)^T \hat{\alpha}_0 (\mathbf{x} - \mathbf{x}_0) \}] . \quad (3)$$

For Equation 2 we have a second derivative term:

$$\begin{aligned} \sum_i \frac{-\hbar^2}{2m_i} \frac{\partial^2}{\partial x_i^2} \Psi(\mathbf{x}, 0) &= \sum_i \frac{-i\hbar}{2m_i} \frac{\partial}{\partial x_i} \{ [\mathbf{p}_0]_i + [\hat{\alpha}_0 (\mathbf{x} - \mathbf{x}_0)]_i + [(\mathbf{x} - \mathbf{x}_0)^T \hat{\alpha}_0]_i \} \Psi(\mathbf{x}, 0) \\ &= \{ -i\hbar \text{Tr}[\hat{\alpha}_0 \hat{m}^{-1}] + \frac{1}{2} \mathbf{p}_0^T \hat{m}^{-1} \mathbf{p}_0 + \mathbf{p}_0^T \hat{m}^{-1} \hat{\alpha}_0 (\mathbf{x} - \mathbf{x}_0) + \\ &\quad (\mathbf{x} - \mathbf{x}_0)^T \hat{\alpha}_0 \hat{m}^{-1} \mathbf{p}_0 + 2 (\mathbf{x} - \mathbf{x}_0)^T \hat{\alpha}_0 \hat{m}^{-1} \hat{\alpha}_0 (\mathbf{x} - \mathbf{x}_0) \} \Psi(\mathbf{x}, 0) , \end{aligned} \quad (4)$$

and terms from the potential expanded to the quadratic term around \mathbf{x}_0 (as in Ref. 1):

$$V(\mathbf{x}) \approx V(\mathbf{x}_0) + \mathbf{V}'(\mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T \hat{V}''(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) . \quad (5)$$

All terms in Eqs. 4 and 5 are of order $\varepsilon\hbar$. Thus they can be collected and returned to the exponential form (accurate up to order ε):

$$\begin{aligned} \Psi(\mathbf{x}, \varepsilon\hbar) \approx \exp \left[\frac{i}{\hbar} \left\{ \gamma_0 + \varepsilon\hbar(i\hbar \text{Tr}[\hat{\alpha}_0 \hat{m}^{-1}] + \frac{1}{2} \mathbf{p}_0 \hat{m}^{-1} \mathbf{p}_0 - V(\mathbf{x}_0)) \right. \right. \\ \left. \left. + \{ \mathbf{p}_0 - \varepsilon\hbar \mathbf{V}'(\mathbf{x}_0) \}^T (\mathbf{x} - \mathbf{x}_0 - \varepsilon\hbar \hat{m}^{-1} \mathbf{p}_0) \right. \right. \\ \left. \left. + (\mathbf{x} - \mathbf{x}_0 - \varepsilon\hbar \hat{m}^{-1} \mathbf{p}_0)^T \left\{ \hat{\alpha}_0 - \varepsilon\hbar \left(2\hat{\alpha}_0 \hat{m}^{-1} \hat{\alpha}_0 + \frac{\hat{V}''}{2} \right) \right\} (\mathbf{x} - \mathbf{x}_0 - \varepsilon\hbar \hat{m}^{-1} \mathbf{p}_0) \right\} + O(\varepsilon^2) \right] . \end{aligned} \quad (6)$$

By defining updated Gaussian variables we now have equations of motion which are accurate up to first order in the time step $dt = \varepsilon\hbar$:

$$\dot{\mathbf{x}}_0 = \hat{m}^{-1} \mathbf{p}_0 , \quad (7)$$

$$\dot{\mathbf{p}}_0 = -\mathbf{V}'(\mathbf{x}_0) , \quad (8)$$

$$\dot{\hat{\alpha}}_0 = -2\hat{\alpha}_0 \hat{m}^{-1} \hat{\alpha}_0 - \frac{\hat{V}''}{2} , \quad (9)$$

$$\dot{\gamma}_0 = i\hbar \text{Tr}[\hat{\alpha}_0 \hat{m}^{-1}] + \frac{1}{2} \mathbf{p}_0 \hat{m}^{-1} \mathbf{p}_0 - V(\mathbf{x}_0) . \quad (10)$$

This result is exactly that of Heller's multidimensional thawed Gaussian wavepacket dynamics.[1]

II. GENERALIZATION TO MULTI-STATE SYSTEM

We again begin with the time dependent Schrödinger Equation. This time our wavefunction is a vector, and the Hamiltonian is a matrix, in electronic state space (defined for a specific geometry \mathbf{x}):

$$|\Psi(\mathbf{x}, t)\rangle = e^{-i\varepsilon \hat{H}(\mathbf{x})} e^{-i\varepsilon \hat{H}(\mathbf{x})} \dots e^{-i\varepsilon \hat{H}(\mathbf{x})} e^{-i\varepsilon \hat{H}(\mathbf{x})} |\Psi(\mathbf{x}, 0)\rangle . \quad (11)$$

The initial wavefunction can undergo a *Born-Oppenheimer expansion*, where for each \mathbf{x} the wavefunction is expanded in a basis of eigenstates of $\hat{V}(\mathbf{x})$:

$$|\Psi(\mathbf{x}, 0)\rangle = \sum_n |n[\mathbf{x}]\rangle \langle n[\mathbf{x}] | \Psi(\mathbf{x}, 0) \rangle . \quad (12)$$

However, here we will take a different approach. For all \mathbf{x} we will expand in the eigenstates of $\hat{V}(\mathbf{x})$ at $\mathbf{x} = \mathbf{x}_0$, where \mathbf{x}_0 is the center of the Gaussian wavepacket $\langle n[\mathbf{x}] | \Psi(\mathbf{x}, 0) \rangle$:

$$|\Psi(\mathbf{x}, 0)\rangle = \sum_n |n[\mathbf{x}_0]\rangle \langle n[\mathbf{x}_0] | \Psi(\mathbf{x}, 0) \rangle . \quad (13)$$

While this choice is formally legal, $\sum_n |n[\mathbf{x}_0]\rangle\langle n[\mathbf{x}_0]| = 1$, it may seem a strange choice. However, it is fully consistent with the trajectory based branching scheme that will ultimately be used to solve this system of equations, and gives the correct receipt for “hopping” trajectories’ boundary conditions. We will consider the wavepacket at time $\hbar\varepsilon$ on state m , which is projected onto the basis of eigenstates of $\hat{V}(\mathbf{x}_1)$, where \mathbf{x}_1 is the center of the wavepacket at time $\hbar\varepsilon$. Taking only the first ε step in Eq. 11, and expanding the exponential to first order in ε we have:

$$\langle m[\mathbf{x}_1]|\Psi(\mathbf{x}, \hbar\varepsilon)\rangle \approx \sum_n \langle m[\mathbf{x}_1]|\{\hat{I} - i\varepsilon\hat{K}(\mathbf{x}) - i\varepsilon\hat{V}(\mathbf{x})\}|n[\mathbf{x}_0]\rangle\langle n[\mathbf{x}_0]|\Psi(\mathbf{x}, 0)\rangle. \quad (14)$$

Eq. 14 describes the wavepacket at time $\hbar\varepsilon$ on electronic surface m , which has contributions from wavepackets at time 0 on all surfaces (n). The eigenstate $|m[\mathbf{x}_1]\rangle$ can be projected in the basis of $|\chi[\mathbf{x}_0]\rangle$ eigenstates (up to the first order in ε):

$$\begin{aligned} |m\rangle &= \sum_l |l\rangle\langle l|m\rangle = \sum_l \{\delta_{l,m} + \langle l|\nabla_{\mathbf{x}_1}m\rangle \cdot (\mathbf{x}_1 - \mathbf{x}_0)\}|l\rangle + \dots \\ &\approx |m\rangle + \varepsilon\hbar \sum_l \mathbf{d}_{l,m} \cdot \hat{m}^{-1}\mathbf{p}_0|l\rangle + O(\varepsilon^2). \end{aligned} \quad (15)$$

We have assumed that the difference $\mathbf{x}_1 - \mathbf{x}_0$ is proportional to $\hbar\varepsilon$ and that we have constant momentum over that time step. \mathbf{p}_0 is the momentum of $\langle n[\mathbf{x}_0]|\Psi(\mathbf{x}, 0)\rangle$. From this point we will drop the $[\mathbf{x}_0]$ label and assume, unless otherwise labeled, that our electronic basis states are eigenstates of $\hat{V}(\mathbf{x}_0)$.

If we insert Eq. 15 into Eq. 14:

$$\begin{aligned} \langle m[\mathbf{x}_1]|\Psi(\mathbf{x}, \hbar\varepsilon)\rangle &\approx \sum_n \left[\langle m| + \hbar\varepsilon \sum_l \mathbf{d}_{l,m} \cdot \hat{m}^{-1}\mathbf{p}_0\langle l| \right] \times \\ &\quad \{1 - i\varepsilon(\hat{K}(\mathbf{x}) + \hat{V}(\mathbf{x}))\}|n\rangle\langle n|\Psi(\mathbf{x}, 0)\rangle, \end{aligned} \quad (16)$$

and separate out the term $m = n$ in the sum, keep only terms up to $O(\varepsilon)$, we have:

$$\begin{aligned} \langle m[\mathbf{x}_1]|\Psi(\mathbf{x}, \hbar\varepsilon)\rangle &\approx \\ &\langle m|\{\hat{I} - i\varepsilon\hat{K}(\mathbf{x}) - i\varepsilon\hat{V}(\mathbf{x})\}|m\rangle\langle m|\Psi(\mathbf{x}, 0)\rangle \\ &- i\varepsilon \sum_{n \neq m} \left[i\hbar\mathbf{d}_{n,m} \cdot \hat{m}^{-1}\mathbf{p}_0 + \langle m|\{\hat{K}(\mathbf{x}) + \hat{V}(\mathbf{x})\}|n\rangle \right] \langle n|\Psi(\mathbf{x}, 0)\rangle. \end{aligned} \quad (17)$$

Here we have used the fact that $\mathbf{d}_{m,m} = 0$ and $|m/n\rangle$ is an eigenstate in order to reduce terms.

We now expand the potential energy matrix operator around \mathbf{x}_0 :

$$\begin{aligned} \hat{V}(\mathbf{x}) &\approx \hat{V}(\mathbf{x}_0) + \hat{\mathbf{V}}'(\mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T \hat{\mathbf{V}}''(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) \\ &= \sum_{\alpha,\beta} \left\{ V_{\alpha,\beta}(\mathbf{x}_0) + \mathbf{V}'_{\alpha,\beta}(\mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T \mathbf{V}''_{\alpha,\beta}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) \right\} |\alpha\rangle\langle\beta|, \end{aligned} \quad (18)$$

where the α/β electronic basis set is \mathbf{x}_0 independent, *e.g* the atomic orbital basis set in realistic calculations, or the diabatic basis set in most model problems. Inserting Eq. 18 into 17 leads to:

$$\begin{aligned} \langle m[\mathbf{x}_1]|\Psi(\mathbf{x}, \hbar\varepsilon)\rangle \approx & \quad (19) \\ \{1 - i\varepsilon[\sum_i \frac{-\hbar^2}{2M_i} \frac{\partial^2}{\partial x_i^2} + E_m(\mathbf{x}_0) - \mathbf{F}_m(\mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0)^T \tilde{\mathbf{H}}_m(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)]\} \langle m|\Psi(\mathbf{x}, 0)\rangle \\ - i\varepsilon \sum_{n \neq m} \left[i\hbar \mathbf{d}_{n,m} \cdot \hat{m}^{-1} \mathbf{p}_0 + \mathbf{V}'_{m,n}(\mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^T \mathbf{V}''_{m,n}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) \right] \langle n|\Psi(\mathbf{x}, 0)\rangle . \end{aligned}$$

Now we consider the form

$$\langle n|\Psi(\mathbf{x}, 0)\rangle \equiv \exp\left[\frac{i}{\hbar}\{\gamma_0 + \mathbf{p}_0^T(\mathbf{x} - \mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0)^T \hat{\alpha}_0(\mathbf{x} - \mathbf{x}_0)\}\right] \times N_n . \quad (20)$$

The first line ($m = n$ case) of Eq. 19 is nearly identical to Eq. 2 with some minor differences. First, the real weight of the wavepacket (N_n) arises since the initial state need not be pure. Second while the first expansion term is the force vector $\langle m|\hat{\mathbf{V}}'(\mathbf{x}_0)|m\rangle = \frac{\partial}{\partial \mathbf{x}} E_m(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_0} = -\mathbf{F}_m(\mathbf{x}_0)$ due to the Hellman Feynman theorem, the second expansion term is not the true Hessian matrix of the potential energy surface ($\tilde{\mathbf{H}}_m(\mathbf{x}_0) \neq \frac{\partial^2}{\partial \mathbf{x}^2} E_m(\mathbf{x})|_{\mathbf{x}=\mathbf{x}_0}$), because we first expanded in the basis which is \mathbf{x}_0 invariant, then rotated into the eigenbasis of $\hat{V}(\mathbf{x}_0)$. Thus we have:

$$\sum_{\alpha, \beta} \langle m|\alpha\rangle \mathbf{V}''_{\alpha, \beta}(\mathbf{x}_0) \langle \beta|m\rangle \equiv \tilde{\mathbf{H}}_m(\mathbf{x}_0) . \quad (21)$$

Finally there is an additional zeroth order in $(\mathbf{x} - \mathbf{x}_0)$ term, coming from the $K_{m,m}(\mathbf{x}_0)$.

$$K_{m,m}(\mathbf{x}_0) \langle m|\Psi(\mathbf{x}, 0)\rangle = -\frac{\hbar^2}{2m_i} \frac{\partial^2}{\partial x_i^2} \langle m|\Psi(\mathbf{x}, 0)\rangle . \quad (22)$$

Thus this first line tells us that the diagonal term is just Heller's *Thawed Gaussian* wavepacket dynamics, but with the second derivative matrix given by Eq. 21.

Now we turn our attention to the $m \neq n$ terms $\mathbf{V}'_{m,n}(\mathbf{x}_0)$ and $\hat{\mathbf{V}}''_{m,n}(\mathbf{x}_0)$. The first derivative term is related to the non-adiabatic coupling vectors through the Hellman-Feynman theorem:

$$\mathbf{V}'_{m,n}(\mathbf{x}_0) = \langle m[\mathbf{x}]|\hat{\mathbf{V}}'(\mathbf{x})|n[\mathbf{x}]\rangle \Big|_{\mathbf{x}=\mathbf{x}_0} = \mathbf{d}_{n,m}^T [E_m(\mathbf{x}_0) - E_n(\mathbf{x}_0)] . \quad (23)$$

The second derivative is similar to the diagonal case:

$$\mathbf{V}''_{m,n}(\mathbf{x}_0) = \sum_{\alpha, \beta} \langle m|\alpha\rangle \mathbf{V}''_{\alpha, \beta}(\mathbf{x}_0) \langle \beta|n\rangle . \quad (24)$$

Collecting all the terms, up to $O(\varepsilon)$, and defining $D_{n,m} \equiv \mathbf{d}_{n,m}^T \hat{m}^{-1} \mathbf{p}_0$ we have:

$$\begin{aligned} \langle m[\mathbf{x}_1]|\Psi(\mathbf{x}, \hbar\varepsilon)\rangle \approx & \dots + \hbar\varepsilon \sum_{n \neq m} D_{n,m} \left[1 + \frac{i}{\hbar} D_{n,m}^{-1} [E_n(\mathbf{x}_0) - E_m(\mathbf{x}_0)] \mathbf{d}_{n,m}^T (\mathbf{x} - \mathbf{x}_0) \right. \\ & \left. - \frac{i}{\hbar} \frac{1}{2} (\mathbf{x} - \mathbf{x}_0)^T D_{n,m}^{-1} \hat{V}''_{m,n}(\mathbf{x}_0) (\mathbf{x} - \mathbf{x}_0) \right] \langle n|\Psi(\mathbf{x}, 0)\rangle . \end{aligned} \quad (25)$$

We can reduce the expression to :

$$\langle m[\mathbf{x}_1]|\Psi(\mathbf{x}, \hbar\varepsilon)\rangle \approx \equiv \dots + \hbar\varepsilon \sum_{n \neq m} D_{n,m} \left[1 + \frac{i}{\hbar} \mathbf{\Delta P}_{m,n}^T (\mathbf{x} - \mathbf{x}_0) + \frac{i}{\hbar} (\mathbf{x} - \mathbf{x}_0)^T \Delta \hat{\alpha}_{m,n} (\mathbf{x} - \mathbf{x}_0) \right]. \quad (26)$$

The term inside the bracket can be exponentiated (assuming all terms are small):

$$\begin{aligned} & \exp \left\{ \frac{i}{\hbar} \mathbf{\Delta P}_{m,n}^T (\mathbf{x} - \mathbf{x}_0) + \frac{i}{\hbar} (\mathbf{x} - \mathbf{x}_0)^T \Delta \hat{\alpha}_{m,n} (\mathbf{x} - \mathbf{x}_0) \right\} \\ & \approx \left[1 + \frac{i}{\hbar} \mathbf{\Delta P}_{m,n}^T (\mathbf{x} - \mathbf{x}_0) + \frac{i}{\hbar} (\mathbf{x} - \mathbf{x}_0)^T \Delta \hat{\alpha}_{m,n} (\mathbf{x} - \mathbf{x}_0) \right], \\ & \text{where } \Delta \hat{\alpha}_{m,n} = -\frac{1}{2} \frac{\hat{V}_{m,n}''(\mathbf{x}_0)}{\mathbf{d}_{n,m}^T \hat{m}^{-1} \mathbf{p}_0}, \\ & \text{and } \mathbf{\Delta P}_{m,n} = \frac{[E_n(\mathbf{x}_0) - E_m(\mathbf{x}_0)]}{\mathbf{d}_{n,m}^T \hat{m}^{-1} \mathbf{p}_0} \mathbf{d}_{n,m}^T. \end{aligned} \quad (27)$$

This condition for the shift in momentum on hop is the same as previously derived. It conserves energy approximately (exactly in infinitely high momentum limit). To ensure that all trajectories conserve energy exactly for all momenta, we make the approximate transformation $\mathbf{d}_{n,m}^T \hat{m}^{-1} \mathbf{p}_0 = \mathbf{d}_{n,m}^T \hat{m}^{-1} \{\mathbf{p}_0 + \mathbf{p}_1\}/2 + \mathbf{d}_{n,m}^T \hat{m}^{-1} \{\mathbf{p}_0 - \mathbf{p}_1\}/2 \approx \mathbf{d}_{n,m}^T \hat{m}^{-1} \{\mathbf{p}_0 + \mathbf{p}_1\}/2 \times \exp \left[\frac{\mathbf{d}_{n,m}^T \{\mathbf{p}_0 - \mathbf{p}_1\}}{\mathbf{d}_{n,m}^T \{\mathbf{p}_0 + \mathbf{p}_1\}} \right]$. With this consideration, and assuming $\mathbf{d}_{n,m}^T \{\mathbf{p}_0 - \mathbf{p}_1\} \ll \mathbf{d}_{n,m}^T \{\mathbf{p}_0 + \mathbf{p}_1\}$ our final result for the ($m \neq n$) case:

$$\begin{aligned} \langle m[\mathbf{x}_1]|\Psi(\mathbf{x}, \hbar\varepsilon)\rangle & \approx \dots + \sum_{n \neq m} \frac{1}{2} \mathbf{d}_{n,m}^T \hat{m}^{-1} \{\mathbf{p}_0 + \mathbf{p}_1\} \times \exp \left[\frac{\mathbf{d}_{n,m}^T \{\mathbf{p}_0 - \mathbf{p}_1\}}{\mathbf{d}_{n,m}^T \{\mathbf{p}_0 + \mathbf{p}_1\}} \right] \times N_n \\ & \times \exp \left[\frac{i}{\hbar} \{ \gamma_0 + \mathbf{p}_1^T (\mathbf{x} - \mathbf{x}_0) + (\mathbf{x} - \mathbf{x}_0)^T \hat{\alpha}_1 (\mathbf{x} - \mathbf{x}_0) \} \right], \\ & \text{where } \mathbf{p}_1 = \mathbf{p}_0 + \frac{[E_n(\mathbf{x}_0) - E_m(\mathbf{x}_0)]}{\mathbf{d}_{n,m}^T \hat{m}^{-1} \{\frac{\mathbf{p}_0 + \mathbf{p}_1}{2}\}} \mathbf{d}_{n,m}^T, \\ & \text{and } \hat{\alpha}_1 = \hat{\alpha}_0 + \frac{\hat{V}_{m,n}''}{\mathbf{d}_{n,m}^T \hat{m}^{-1} \{\mathbf{p}_0 + \mathbf{p}_1\}}. \end{aligned} \quad (28)$$

III. WAVEPACKET RECONSTRUCTION

For Coupled Propagation and GWP Consolidation (See Section IV) we need to calculate overlap of two normalized Gaussians. We seek to define a single Gaussian which closely approximates two separate but similar Gaussians:

$$N_G e^{i\gamma_G} |G(\mathbf{x}_G, \mathbf{p}_G, \hat{\alpha}_G)\rangle \approx N_1 e^{i\gamma_1} |g_1(\mathbf{x}_1, \mathbf{p}_1, \hat{\alpha}_1)\rangle + N_2 e^{i\gamma_2} |g_2(\mathbf{x}_2, \mathbf{p}_2, \hat{\alpha}_2)\rangle \equiv |\psi\rangle \quad (29)$$

where G , g_1 and g_2 are normalized unphased complex Gaussians. In the limit that the superposition $|\psi\rangle$ is indeed a Gaussian then the mapping is exact:

$$N_G = |\langle\psi|\psi\rangle|^{\frac{1}{2}}, \quad (30)$$

$$\mathbf{x}_G = \frac{\langle\psi|\mathbf{x}|\psi\rangle}{\langle\psi|\psi\rangle}, \quad (31)$$

$$\mathbf{p}_G = -i \frac{\langle\psi|\partial_{\mathbf{x}}|\psi\rangle}{\langle\psi|\psi\rangle}, \quad (32)$$

$$\text{Im}[\hat{\alpha}_G] = \frac{1}{4} \left[\frac{\langle\psi|(\mathbf{x} - \mathbf{x}_G)^2|\psi\rangle}{\langle\psi|\psi\rangle} \right]^{-1}, \quad (33)$$

$$\frac{\langle\psi|(-i\partial_{\mathbf{x}} - \mathbf{p}_G)^2|\psi\rangle}{\langle\psi|\psi\rangle} - \text{Im}[\hat{\alpha}_G] = \text{Re}[\hat{\alpha}_G] \text{Im}[\hat{\alpha}_G]^{-1} \text{Re}[\hat{\alpha}_G]. \quad (34)$$

Equation 34 can be solved using the Geometric mean for positive definite matrices $\text{Im}[\hat{\alpha}_G]$ and $\frac{\langle\psi|(-i\partial_{\mathbf{x}} - \mathbf{p}_G)^2|\psi\rangle}{\langle\psi|\psi\rangle} - \text{Im}[\hat{\alpha}_G]$:

$$\text{Re}[\hat{\alpha}_G] = \text{Im}[\hat{\alpha}_G]^{\frac{1}{2}} \left[\text{Im}[\hat{\alpha}_G]^{-\frac{1}{2}} \left\{ \frac{\langle\psi|(-i\partial_{\mathbf{x}} - \mathbf{p}_G)^2|\psi\rangle}{\langle\psi|\psi\rangle} - \text{Im}[\hat{\alpha}_G] \right\} \text{Im}[\hat{\alpha}_G]^{-\frac{1}{2}} \right]^{\frac{1}{2}} \text{Im}[\hat{\alpha}_G]^{\frac{1}{2}}. \quad (35)$$

These expectation values and overlap of a superposition of multivariate Gaussians can be calculated analytically. Finally the phase can be found by maximizing the overlap of $\langle\psi|N_G e^{i\gamma_G} |G(\mathbf{x}_G, \mathbf{p}_G, \hat{\alpha}_G)\rangle$, under the constraint of Equation 30:

$$e^{i\gamma_G} = \frac{\langle G(\mathbf{x}_G, \mathbf{p}_G, \hat{\alpha}_G) | \psi \rangle}{|\langle \psi | G(\mathbf{x}_G, \mathbf{p}_G, \hat{\alpha}_G) \rangle|}. \quad (36)$$

The expectation values of the superposition are given as a sum over combinations of the Gaussians:

$$\langle\psi|O|\psi\rangle = \sum_{i,j \in 1,2} \langle g_j | O | g_i \rangle N_i N_j \exp[i\{\gamma_i - \gamma_j\}]. \quad (37)$$

Through Equations 31-32 the dynamics of the Coupled GWP depends strongly on the Thawed Gaussian Approximation, and the particular value of $\hat{\alpha}$. As the Coupled GWPs separate and the approximation of Equations 30-36 become less valid, the dependence of Equations 31-32 on $\hat{\alpha}$ can lead to unstable dynamics. This is particularly true when the value of $\text{Im}[\hat{\alpha}]$ becomes small (a very wide wavepacket). One valuable feature of the Thawed Gaussian Approximation is that the dynamics are fully classical, and independent of the phase and width. In the same spirit, here we seek to add further approximations to the dynamics which will add stability, with minimal sacrifice of accuracy. We note that typically we will break coupling when $|\langle g_1 | g_2 \rangle|$ is much less than unity. The quantity $\langle g_1 | g_2 \rangle$ appears in the evaluation of all Equations 30-36.

$$\begin{aligned} \langle g_1 | g_2 \rangle = & (\text{Det}[\frac{2}{\pi} \hat{\alpha}_2])^{\frac{1}{4}} (\text{Det}[\frac{2}{\pi} \hat{\alpha}_1])^{\frac{1}{4}} \int d\mathbf{x} \exp[i\{(\mathbf{x} - \mathbf{x}_2) \hat{\alpha}_2 (\mathbf{x} - \mathbf{x}_2) - (\mathbf{x} - \mathbf{x}_1) \hat{\alpha}_1^* (\mathbf{x} - \mathbf{x}_1)\}] \\ & \times \exp[i\{\mathbf{p}_1 \mathbf{x}_1 - \mathbf{p}_2 \mathbf{x}_2\}] \times \exp[i\{\mathbf{p}_2 - \mathbf{p}_1\} \mathbf{x}]. \end{aligned} \quad (38)$$

To stabilize the dynamics we make approximations when finding \mathbf{x}_G and \mathbf{p}_G . For high momentum and close narrow wavepackets:

$$(\text{Det}[\frac{2}{\pi}\hat{\alpha}_2])^{\frac{1}{4}}(\text{Det}[\frac{2}{\pi}\hat{\alpha}_1])^{\frac{1}{4}}\exp[i\{(\mathbf{x} - \mathbf{x}_2)\hat{\alpha}_2(\mathbf{x} - \mathbf{x}_2) - (\mathbf{x} - \mathbf{x}_1)\hat{\alpha}_1^*(\mathbf{x} - \mathbf{x}_1)\}] \approx \delta(\mathbf{x} - [\frac{\mathbf{x}_1 + \mathbf{x}_2}{2}]) ,$$

which leads to:

$$\langle g_1|g_2\rangle \approx \exp[i\{\frac{\mathbf{p}_2 + \mathbf{p}_1}{2}\}\{\mathbf{x}_1 - \mathbf{x}_2\}] . \quad (39)$$

Use of Equation 39 leads to:

$$\langle g_1|\mathbf{x}|g_2\rangle = -i\partial_{\mathbf{p}_2}\langle g_1|g_2\rangle + \langle g_1|\mathbf{x}_1|g_2\rangle \approx \frac{\mathbf{x}_1 + \mathbf{x}_2}{2}\exp[i\{\frac{\mathbf{p}_2 + \mathbf{p}_1}{2}\}\{\mathbf{x}_1 - \mathbf{x}_2\}] , \quad (40)$$

$$\begin{aligned} -i\langle g_1|\partial_{\mathbf{x}}|g_2\rangle &= -\frac{i}{2}[\langle g_1|\partial_{\mathbf{x}}g_2\rangle - \langle \partial_{\mathbf{x}}g_1|g_2\rangle] = \frac{\mathbf{p}_1 + \mathbf{p}_2}{2}\langle g_1|g_2\rangle + \hat{\alpha}_2\langle g_1|\mathbf{x} - \mathbf{x}_2|g_2\rangle + \hat{\alpha}_1^*\langle g_1|\mathbf{x} - \mathbf{x}_1|g_2\rangle \\ &\approx \frac{\mathbf{p}_1 + \mathbf{p}_2}{2}\exp[i\{\frac{\mathbf{p}_2 + \mathbf{p}_1}{2}\}\{\mathbf{x}_1 - \mathbf{x}_2\}] . \end{aligned} \quad (41)$$

Equations 40 and 41 are used to in the calculations shown in the main text. In the calculation of Equation 41 we discard the terms which are proportional to $\hat{\alpha}$. This is consistent with the approximation leading to Equation 39, that momentum is high and wave packets are narrow ($\frac{\mathbf{p}_1 + \mathbf{p}_2}{2} \gg \frac{1}{2}\{\mathbf{x}_1 - \mathbf{x}_2\}\{\hat{\alpha}_2 - \hat{\alpha}_1^*\}$). For the calculation of $\hat{\alpha}_G$ the full set of Equations 30-34 are used, but with the GWPs following the dynamics guided by Equations 40 and 41. Using Equation 39 in the evaluation of N_G and γ_G (Equations 30 and 36) provides similar results to using the full $\langle g_1|g_2\rangle$ calculation for the models considered in the main text.

IV. SUMMARY OF THE ALGORITHM

Initial Propagation:

Step 1: Initialize Gaussian wavepacket (GWP) with desired parameters, on state m .

Step 2: Propagate the GWP forward in time using Eq. 5 from the main text.

Step 3: If outside region of non-adiabatic coupling (NAC) repeat step 2. If GWP reaches a region of significant NAC, $\mathbf{d}_{n,m}^T \frac{\hat{m}^{-1}\{\mathbf{p}_0 + \mathbf{p}_1\}/2}{|\hat{m}^{-1}\{\mathbf{p}_0 + \mathbf{p}_1\}/2|} > D_{min}$ (a user set threshold), then generate new wavepacket on state n using Eq. 6 from main text and energy conserving change in momentum. If energy cannot be conserved, generation is not allowed. Establish a connection between these GWPs on m and n , and begin Coupled Propagation.

Coupled Propagation:

Step 4: Propagate all GWP forward in time using Eq. 5 from main text.

Step 5: Calculate overlap of normalized GWP with it's connections normalized "hopped" GWP (with width and momentum shift described above). If the amplitude of the overlap is greater than an accuracy controlling threshold, O_{min} , reconstruct new wavepacket as described in Section III . If the amplitude of the overlap is lower than O_{min} , eliminate connection between these GWPs. Continue with Step 2 for each wavepacket independently.

Step 6: If a GWP leaves the region of significant NAC, $\mathbf{d}_{n,m}^T \frac{\hat{m}^{-1}\{\mathbf{p}_0+\mathbf{p}_1\}/2}{|\hat{m}^{-1}\{\mathbf{p}_0+\mathbf{p}_1\}/2|} < D_{min}$, eliminate connection between these GWPs. Continue with Step 2 for each wavepacket independently.

End of Propagation:

Step 7: Once final (or output) time is reached, the wavefunction is given as a sum over the weighted complex GWPs. Expectation values can be calculated directly from the wavefunction.

Consolidation and Filtering (Optional Consideration):

Periodically, on some predefined number of time step, one could attempt to condense the trajectories by using the same reconstruction method as used in coupled propagation. Additionally after such a consolidation step one could discard trajectories which are insignificant (by real weight).

Step A: Remove connection between all coupled trajectories.

Step B: Loop through PES,

Step C: Find the largest weight GWP on the PES, this GWP is the initial value of the reconstructed GWP (GWP-New).

Step D: Loop through all available GWPs on the PES. If the normalized GWP has overlap with the normalized GWP-New is higher than O_{min} , then add the GWP to GWP-New (by reconstruction). Continue to update GWP-New until fully looped through the GWP's. Repeat Loop until no new GWPs are added to GWP-New.

Step E: Repeat Step C, until all GWPs have been added to new GWPs.

Step F: Return to Step C for next PES, continuing until all PES are checked.

Step G: Once all new, consolidated GWPs are generated, calculate average weight of GWPs. For all GWPs, if the weight is less than a threshold percentage of the average weight (we have used 3%) then discard the GWP from further propagation.

Step H: Procedure (starting from Step B) can be repeated, until no further change in the GWPs occurs, to attempt further consolidation.

Step I: For all GWPs on different PES: If the overlap of the normalized GWPs is greater than a threshold, they are re-connected and undergo Coupled Propagation (Step 4), else if no “partner” is found then the GWP undergoes independent propagation (Step 2).

Monte-Carlo Sampling of many branch GWPs (Optional Consideration):

If a GWP is the result of many branches one expects its contribution to be small, with many similar GWPs contributing. Thus for branches of higher order than a defined number B_{max} , one may choose to sample the branching by Monte-Carlo rather than explicitly propagating both branches.

Step 5/6-a: If connection between GWPs is ended, and the GWPs have branched more than B_{max} times, choose one trajectory to propagate by random number generation (Monte Carlo), based on the weights of the GWP. The propagated GWP will have new weight which is equal to the sum of the weights of the two GWPs.

Step 5/6-b: New GWPs generated by the Consolidation and Filtering procedure will be considered as having branched zero times.

Miscellaneous considerations to limit branches (Optional Considerations):

6-M: (Modified Step 6) If GWP leave the region of significant non-adiabatic coupling (NAC), $\mathbf{d}_{n,m}^T \frac{\hat{m}^{-1}\{\mathbf{p}_0+\mathbf{p}_1\}/2}{|\hat{m}^{-1}\{\mathbf{p}_0+\mathbf{p}_1\}/2|} < D_{min}$, eliminate connection in single direction, allowing GWP which is still inside NAC region to add to the GWP which has left, but not the reverse. This can help prevent the generation of many small GWPs as the coupled GWPs leave the region of NAC at slightly different times.

Step CP-M: (Modification to Coupled Propagation) Only check whether to break connection after GWPs are connected for a finite time ($t_{min} = 10$ a.u.). This helps prevent generation

of many small wave packets in difficult, highly-chaoitic, regions or near edges of NAC region.

Step 3-M: (Modification to Step 3) One can define a region of significant NAC where the Massey parameter[2] for each GWP:

$$\zeta_0 = \left| \frac{\mathbf{d}_{n,m}^T \hat{m}^{-1} \mathbf{p}_0}{E_n - E_m} \right| / \zeta_1 = \left| \frac{\mathbf{d}_{m,n}^T \hat{m}^{-1} \mathbf{p}_1}{E_m - E_n} \right| > \zeta_{min} . \quad (42)$$

We use $\zeta_{min} = 1E - 3$. One may also limit the NAC region to areas where $\frac{\mathbf{d}_{n,m}^T \{\mathbf{p}_0 - \mathbf{p}_1\}}{\mathbf{d}_{n,m}^T \{\mathbf{p}_0 + \mathbf{p}_1\}} < \Delta_{max}$ (we use $\Delta_{max} = 1.5$). When this ratio is large rapidly oscillating phase differences between generated GWPs is expected to cancel out.

Note: With all the optional considerations, there is a threshold parameter ($O_{min}, D_{min}, B_{max}, t_{min}, \zeta_{min}, \Delta_{max}$) which can be used to tune accuracy vs efficiency. Convergence with the thresholds can be checked to determine if the information loss is acceptable.

V. NON-ORTHOGONALITY OF BASIS

During Coupled Propagation, we must take a superposition of two Gaussians which are in non-orthogonal basis states. For two coupled GWPs, which generate two “hopped” GWPs we have:

$$\begin{aligned} |\Psi\rangle &= g_1|1[\mathbf{x}_1]\rangle + g_{1,2}|2[\mathbf{x}_1]\rangle + g_{2,1}|1[\mathbf{x}_2]\rangle + g_2|2[\mathbf{x}_2]\rangle \\ &\approx G_1|1[\mathbf{x}_1]\rangle + G_2|2[\mathbf{x}_2]\rangle . \end{aligned} \quad (43)$$

Here, $g_{1(2)}$ is the GWP initially on PES 1 (2) which stays on PES 1(2), and $g_{1,2(2,1)}$ is the GWP which is initially on PES 1 (2) and hops to PES 2(1). The GWPs $g_{1,2}$ and $g_{2,1}$ must be rotated from the electronic states $|2[\mathbf{x}_1]\rangle$ and $|1[\mathbf{x}_2]\rangle$ to $|2[\mathbf{x}_2]\rangle$ and $|1[\mathbf{x}_1]\rangle$ respectively. Projection of $|2[\mathbf{x}_1]\rangle$ onto the orthogonal basis $|2[\mathbf{x}_2]\rangle + |1[\mathbf{x}_2]\rangle$ will result in the wavepacket being “duplicated”:

$$g_{1,2}|2[\mathbf{x}_1]\rangle = g_{1,2} \cos\theta |2[\mathbf{x}_2]\rangle + g_{1,2} \sin\theta |1[\mathbf{x}_2]\rangle , \quad (44)$$

and similarly, we have:

$$g_{2,1}|1[\mathbf{x}_2]\rangle = g_{2,1} \cos\theta |1[\mathbf{x}_1]\rangle + g_{2,1} \sin\theta |2[\mathbf{x}_1]\rangle . \quad (45)$$

where $\cos\theta \equiv \langle 1[\mathbf{x}_2]|1[\mathbf{x}_1]\rangle \equiv \langle 2[\mathbf{x}_2]|2[\mathbf{x}_1]\rangle$ and $\sin\theta \equiv \langle 1[\mathbf{x}_2]|2[\mathbf{x}_1]\rangle \equiv \langle 2[\mathbf{x}_1]|1[\mathbf{x}_2]\rangle$. This step can be repeated, and since by definition $|\sin\theta| < 1$ the infinite series of rotations will converge. Thus

Equation 43 can be re-expressed as:

$$\begin{aligned}
 |\Psi\rangle &= g_1|1[\mathbf{x}_1]\rangle + g_{1,2}|2[\mathbf{x}_1]\rangle + g_{2,1}|1[\mathbf{x}_2]\rangle + g_2|2[\mathbf{x}_2]\rangle \\
 &= \left[g_1 + \frac{1}{\cos\theta}g_{2,1} + \frac{\sin\theta}{\cos\theta}g_{1,2} \right] |1[\mathbf{x}_1]\rangle + \left[g_2 + \frac{1}{\cos\theta}g_{1,2} + \frac{\sin\theta}{\cos\theta}g_{2,1} \right] |2[\mathbf{x}_2]\rangle \\
 &\approx G_1|1[\mathbf{x}_1]\rangle + G_2|2[\mathbf{x}_2]\rangle.
 \end{aligned} \tag{46}$$

Thus, not only does the “hopped” GWP contribute to the reconstructed GWP on the final surface, but it also has a contribution to the re-constructed GWP on the same surface. However this contributions is small for shifts in position $\mathbf{x}_1 - \mathbf{x}_2$ in which the Coupled GWP dynamics works ($\cos\theta \approx 1$). That is, GWPs should branch before this non-orthogonal basis effect becomes relevant. In the calculations presented in the paper, inclusion of this effect does not change results significantly. We present it here for formal completeness, and for possible future importance.

-
- [1] E. J. Heller, The Journal of Chemical Physics **62** (1975).
 [2] H. S. W. Massey, Reports on Progress in Physics **12**, 248 (1949).